OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/853,033

DATE: 08/14/2001
TIME: 12:13:52

Input Set : A:\65691222.app
Output Set: N:\CRF3\08142001\I853033.raw

ENTERED

```
 3 <110> APPLICANT: CHAMBON, PIERRE
 4         METZGER, DANIEL
 6 <120> TITLE OF INVENTION: TRANSGENIC MOUSE FOR TARGETED RECOMBINATION
 7         MEDIATED BY MODIFIED CRE-ER
 9 <130> FILE REFERENCE: 065691/0222
11 <140> CURRENT APPLICATION NUMBER: 09/853,033
12 <141> CURRENT FILING DATE: 2001-05-11
14 <150> PRIOR APPLICATION NUMBER: FR 00/12570
15 <151> PRIOR FILING DATE: 2000-10-03
17 <160> NUMBER OF SEQ ID NOS: 14
19 <170> SOFTWARE: PatentIn Ver. 2.1
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 1788
23 <212> TYPE: DNA
24 <213> ORGANISM: Homo sapiens
26 <220> FEATURE:
27 <221> NAME/KEY: CDS
28 <222> LOCATION: (1)..(1788)
30 <400> SEQUENCE: 1
31 atg acc atg acc ctc cac acc aaa gca tct ggg atg gcc cta ctg cat    48
32 Met Thr Met Thr Leu His Thr Lys Ala Ser Gly Met Ala Leu Leu His
33   1               5                  10                  15
35 cag atc caa ggg aac gag ctg gag ccc ctg aac cgt ccg cag ctc aag    96
36 Gln Ile Gln Gly Asn Glu Leu Glu Pro Leu Asn Arg Pro Gln Leu Lys
37              20                  25                  30
39 atc ccc ctg gag cgg ccc ctg ggc gag gtg tac ctg gac agc agc aag   144
40 Ile Pro Leu Glu Arg Pro Leu Gly Glu Val Tyr Leu Asp Ser Ser Lys
41          35                  40                  45
43 ccc gcc gtg tac aac tac ccc gag ggc gcc gcc tac gag ttc aac gcc   192
44 Pro Ala Val Tyr Asn Tyr Pro Glu Gly Ala Ala Tyr Glu Phe Asn Ala
45      50                  55                  60
47 gcg gcc gcc gcc aac gcg cag gtc tac ggt cag acc ggc ctc ccc tac   240
48 Ala Ala Ala Ala Asn Ala Gln Val Tyr Gly Gln Thr Gly Leu Pro Tyr
49  65                  70                  75                  80
51 ggc ccc ggg tct gag gct gcg gcg ttc ggc tcc aac ggc ctg ggg ggt   288
52 Gly Pro Gly Ser Glu Ala Ala Ala Phe Gly Ser Asn Gly Leu Gly Gly
53                  85                  90                  95
55 ttc ccc cca ctc aac agc gtg tct ccg agc ccg ctg atg cta ctg cac   336
56 Phe Pro Pro Leu Asn Ser Val Ser Pro Ser Pro Leu Met Leu Leu His
57                 100                 105                 110
59 ccg ccg ccg cag ctg tcg cct ttc ctg cag ccc cac ggc cag cag gtg   384
60 Pro Pro Pro Gln Leu Ser Pro Phe Leu Gln Pro His Gly Gln Gln Val
61             115                 120                 125
63 ccc tac tac ctg gag aac gag ccc agc ggc tac acg gtg cgc gag gcc   432
64 Pro Tyr Tyr Leu Glu Asn Glu Pro Ser Gly Tyr Thr Val Arg Glu Ala
65         130                 135                 140
67 ggc ccg ccg gca ttc tac agg cca aat tca gat aat cga cgc cag ggt   480
```

```
68 Gly Pro Pro Ala Phe Tyr Arg Pro Asn Ser Asp Asn Arg Arg Gln Gly
69 145                 150             155                 160
71 ggc aga gaa aga ttg gcc agt acc aat gac aag gga agt atg gct atg    528
72 Gly Arg Glu Arg Leu Ala Ser Thr Asn Asp Lys Gly Ser Met Ala Met
73                 165             170                 175
75 gaa tct gcc aag gag act cgc tac tgt gca gtg tgc aat gac tat gct    576
76 Glu Ser Ala Lys Glu Thr Arg Tyr Cys Ala Val Cys Asn Asp Tyr Ala
77             180             185             190
79 tca ggc tac cat tat gga gtc tgg tcc tgt gag ggc tgc aag gcc ttc    624
80 Ser Gly Tyr His Tyr Gly Val Trp Ser Cys Glu Gly Cys Lys Ala Phe
81         195             200             205
83 ttc aag aga agt att caa gga cat aac gac tat atg tgt cca gcc acc    672
84 Phe Lys Arg Ser Ile Gln Gly His Asn Asp Tyr Met Cys Pro Ala Thr
85     210             215             220
87 aac cag tgc acc att gat aaa aac agg agg aag agc tgc cag gcc tgc    720
88 Asn Gln Cys Thr Ile Asp Lys Asn Arg Arg Lys Ser Cys Gln Ala Cys
89 225             230             235             240
91 cgg ctc cgc aaa tgc tac gaa gtg gga atg atg aaa ggt ggg ata cga    768
92 Arg Leu Arg Lys Cys Tyr Glu Val Gly Met Met Lys Gly Gly Ile Arg
93             245             250             255
95 aaa gac cga aga gga ggg aga atg ttg aaa cac aag cgc cag aga gat    816
96 Lys Asp Arg Arg Gly Gly Arg Met Leu Lys His Lys Arg Gln Arg Asp
97         260             265             270
99 gat ggg gag ggc agg ggt gaa gtg ggg tct gct gga gac atg aga gct    864
100 Asp Gly Glu Gly Arg Gly Glu Val Gly Ser Ala Gly Asp Met Arg Ala
101         275             280             285
103 gcc aac ctt tgg cca agc ccg ctc atg atc aaa cgc tct aag aag aac    912
104 Ala Asn Leu Trp Pro Ser Pro Leu Met Ile Lys Arg Ser Lys Lys Asn
105     290             295             300
107 agc ctg gcc ttg tcc ctg acg gcc gac cag atg gtc agt gcc ttg ttg    960
108 Ser Leu Ala Leu Ser Leu Thr Ala Asp Gln Met Val Ser Ala Leu Leu
109 305             310             315             320
111 gat gct gag ccc ccc ata ctc tat tcc gag tat gat cct acc aga ccc   1008
112 Asp Ala Glu Pro Pro Ile Leu Tyr Ser Glu Tyr Asp Pro Thr Arg Pro
113             325             330             335
115 ttc agt gaa gct tcg atg atg ggc tta ctg acc aac ctg gca gac agg   1056
116 Phe Ser Glu Ala Ser Met Met Gly Leu Leu Thr Asn Leu Ala Asp Arg
117         340             345             350
119 gag ctg gtt cac atg atc aac tgg gcg aag agg gtg cca ggc ttt gtg   1104
120 Glu Leu Val His Met Ile Asn Trp Ala Lys Arg Val Pro Gly Phe Val
121         355             360             365
123 gat ttg acc ctc cat gat cag gtc cac ctt cta gaa tgt gcc tgg cta   1152
124 Asp Leu Thr Leu His Asp Gln Val His Leu Leu Glu Cys Ala Trp Leu
125     370             375             380
127 gag atc ctg atg att ggt ctc gtc tgg cgc tcc atg gag cac cca ggg   1200
128 Glu Ile Leu Met Ile Gly Leu Val Trp Arg Ser Met Glu His Pro Gly
129 385             390             395             400
131 aag cta ctg ttt gct cct aac ttg ctc ttg gac agg aac cag gga aaa   1248
132 Lys Leu Leu Phe Ala Pro Asn Leu Leu Leu Asp Arg Asn Gln Gly Lys
```

Input Set : A:\65691222.app
Output Set: N:\CRF3\08142001\I853033.raw

```
133                     405                   410                   415
135 tgt gta gag ggc atg gtg gag atc ttc gac atg ctg ctg gct aca tca   1296
136 Cys Val Glu Gly Met Val Glu Ile Phe Asp Met Leu Leu Ala Thr Ser
137                 420                   425                   430
139 tct cgg ttc cgc atg atg aat ctg cag gga gag gag ttt gtg tgc ctc   1344
140 Ser Arg Phe Arg Met Met Asn Leu Gln Gly Glu Glu Phe Val Cys Leu
141             435                   440                   445
143 aaa tct att att ttg ctt aat tct gga gtg tac aca ttt ctg tcc agc   1392
144 Lys Ser Ile Ile Leu Leu Asn Ser Gly Val Tyr Thr Phe Leu Ser Ser
145     450                   455                   460
147 acc ctg aag tct ctg gaa gag aag gac cat atc cac cga gtc ctg gac   1440
148 Thr Leu Lys Ser Leu Glu Glu Lys Asp His Ile His Arg Val Leu Asp
149 465                   470                   475                   480
151 aag atc aca gac act ttg atc cac ctg atg gcc aag gca ggc ctg acc   1488
152 Lys Ile Thr Asp Thr Leu Ile His Leu Met Ala Lys Ala Gly Leu Thr
153                 485                   490                   495
155 ctg cag cag cag cac cag cgg ctg gcc cag ctc ctc ctc atc ctc tcc   1536
156 Leu Gln Gln Gln His Gln Arg Leu Ala Gln Leu Leu Leu Ile Leu Ser
157             500                   505                   510
159 cac atc agg cac atg agt aac aaa ggc atg gag cat ctg tac agc atg   1584
160 His Ile Arg His Met Ser Asn Lys Gly Met Glu His Leu Tyr Ser Met
161             515                   520                   525
163 aag tgc aag aac gtg gtg ccc ctc tat gac ctg ctg ctg gag atg ctg   1632
164 Lys Cys Lys Asn Val Val Pro Leu Tyr Asp Leu Leu Leu Glu Met Leu
165     530                   535                   540
167 gac gcc cac cgc cta cat gcg ccc act agc cgt gga ggg gca tcc gtg   1680
168 Asp Ala His Arg Leu His Ala Pro Thr Ser Arg Gly Gly Ala Ser Val
169 545                   550                   555                   560
171 gag gag acg gac caa agc cac ttg gcc act gcg ggc tct act tca tcg   1728
172 Glu Glu Thr Asp Gln Ser His Leu Ala Thr Ala Gly Ser Thr Ser Ser
173                 565                   570                   575
175 cat tcc ttg caa aag tat tac atc acg ggg gag gca gag ggt ttc cct   1776
176 His Ser Leu Gln Lys Tyr Tyr Ile Thr Gly Glu Ala Glu Gly Phe Pro
177             580                   585                   590
179 gcc aca gtc tga                                                    1788
180 Ala Thr Val
181             595
```

184 <210> SEQ ID NO: 2
185 <211> LENGTH: 595
186 <212> TYPE: PRT
187 <213> ORGANISM: Homo sapiens
189 <400> SEQUENCE: 2

```
190 Met Thr Met Thr Leu His Thr Lys Ala Ser Gly Met Ala Leu Leu His
191 1               5                   10                  15
193 Gln Ile Gln Gly Asn Glu Leu Glu Pro Leu Asn Arg Pro Gln Leu Lys
194             20                  25                  30
196 Ile Pro Leu Glu Arg Pro Leu Gly Glu Val Tyr Leu Asp Ser Ser Lys
197         35                  40                  45
199 Pro Ala Val Tyr Asn Tyr Pro Glu Gly Ala Ala Tyr Glu Phe Asn Ala
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/853,033

Input Set : A:\65691222.app
Output Set: N:\CRF3\08142001\I853033.raw

```
200      50                  55                  60
202 Ala Ala Ala Ala Asn Ala Gln Val Tyr Gly Gln Thr Gly Leu Pro Tyr
203  65              70              75                  80
205 Gly Pro Gly Ser Glu Ala Ala Ala Phe Gly Ser Asn Gly Leu Gly Gly
206              85              90                  95
208 Phe Pro Pro Leu Asn Ser Val Ser Pro Ser Pro Leu Met Leu Leu His
209             100             105             110
211 Pro Pro Pro Gln Leu Ser Pro Phe Leu Gln Pro His Gly Gln Gln Val
212         115             120             125
214 Pro Tyr Tyr Leu Glu Asn Glu Pro Ser Gly Tyr Thr Val Arg Glu Ala
215     130             135             140
217 Gly Pro Pro Ala Phe Tyr Arg Pro Asn Ser Asp Asn Arg Arg Gln Gly
218 145             150             155             160
220 Gly Arg Glu Arg Leu Ala Ser Thr Asn Asp Lys Gly Ser Met Ala Met
221             165             170             175
223 Glu Ser Ala Lys Glu Thr Arg Tyr Cys Ala Val Cys Asn Asp Tyr Ala
224             180             185             190
226 Ser Gly Tyr His Tyr Gly Val Trp Ser Cys Glu Gly Cys Lys Ala Phe
227         195             200             205
229 Phe Lys Arg Ser Ile Gln Gly His Asn Asp Tyr Met Cys Pro Ala Thr
230     210             215             220
232 Asn Gln Cys Thr Ile Asp Lys Asn Arg Arg Lys Ser Cys Gln Ala Cys
233 225             230             235             240
235 Arg Leu Arg Lys Cys Tyr Glu Val Gly Met Met Lys Gly Gly Ile Arg
236             245             250             255
238 Lys Asp Arg Arg Gly Gly Arg Met Leu Lys His Lys Arg Gln Arg Asp
239         260             265             270
241 Asp Gly Glu Gly Arg Gly Glu Val Gly Ser Ala Gly Asp Met Arg Ala
242     275             280             285
244 Ala Asn Leu Trp Pro Ser Pro Leu Met Ile Lys Arg Ser Lys Lys Asn
245 290             295             300
247 Ser Leu Ala Leu Ser Leu Thr Ala Asp Gln Met Val Ser Ala Leu Leu
248 305             310             315             320
250 Asp Ala Glu Pro Pro Ile Leu Tyr Ser Glu Tyr Asp Pro Thr Arg Pro
251             325             330             335
253 Phe Ser Glu Ala Ser Met Met Gly Leu Leu Thr Asn Leu Ala Asp Arg
254             340             345             350
256 Glu Leu Val His Met Ile Asn Trp Ala Lys Arg Val Pro Gly Phe Val
257         355             360             365
259 Asp Leu Thr Leu His Asp Gln Val His Leu Leu Glu Cys Ala Trp Leu
260     370             375             380
262 Glu Ile Leu Met Ile Gly Leu Val Trp Arg Ser Met Glu His Pro Gly
263 385             390             395             400
265 Lys Leu Leu Phe Ala Pro Asn Leu Leu Leu Asp Arg Asn Gln Gly Lys
266             405             410             415
268 Cys Val Glu Gly Met Val Glu Ile Phe Asp Met Leu Leu Ala Thr Ser
269             420             425             430
271 Ser Arg Phe Arg Met Met Asn Leu Gln Gly Glu Glu Phe Val Cys Leu
272         435             440             445
```

Input Set : A:\65691222.app
Output Set: N:\CRF3\08142001\I853033.raw

```
274 Lys Ser Ile Ile Leu Leu Asn Ser Gly Val Tyr Thr Phe Leu Ser Ser
275     450                     455                 460
277 Thr Leu Lys Ser Leu Glu Glu Lys Asp His Ile His Arg Val Leu Asp
278 465                     470                 475                 480
280 Lys Ile Thr Asp Thr Leu Ile His Leu Met Ala Lys Ala Gly Leu Thr
281                     485                 490                 495
283 Leu Gln Gln Gln His Gln Arg Leu Ala Gln Leu Leu Leu Ile Leu Ser
284                 500                 505                 510
286 His Ile Arg His Met Ser Asn Lys Gly Met Glu His Leu Tyr Ser Met
287             515                 520                 525
289 Lys Cys Lys Asn Val Val Pro Leu Tyr Asp Leu Leu Leu Glu Met Leu
290     530                 535                 540
292 Asp Ala His Arg Leu His Ala Pro Thr Ser Arg Gly Gly Ala Ser Val
293 545                 550                 555                 560
295 Glu Glu Thr Asp Gln Ser His Leu Ala Thr Ala Gly Ser Thr Ser Ser
296                 565                 570                 575
298 His Ser Leu Gln Lys Tyr Tyr Ile Thr Gly Glu Ala Glu Gly Phe Pro
299             580                 585                 590
301 Ala Thr Val
302         595
```

```
305 <210> SEQ ID NO: 3
306 <211> LENGTH: 1983
307 <212> TYPE: DNA
308 <213> ORGANISM: Artificial sequence
310 <220> FEATURE:
311 <221> NAME/KEY: CDS
312 <222> LOCATION: (1)..(1983)
314 <220> FEATURE:
315 <223> OTHER INFORMATION: Description of Artificial Sequence: Chimeric sequence
316         Homosapiens-Bacteriophage P1
318 <400> SEQUENCE: 3
319 atg tcc aat tta ctg acc gta cac caa aat ttg cct gca tta ccg gtc     48
320 Met Ser Asn Leu Leu Thr Val His Gln Asn Leu Pro Ala Leu Pro Val
321   1               5                   10                  15
323 gat gca acg agt gat gag gtt cgc aag aac ctg atg gac atg ttc agg     96
324 Asp Ala Thr Ser Asp Glu Val Arg Lys Asn Leu Met Asp Met Phe Arg
325               20                  25                  30
327 gat cgc cag gcg ttt tct gag cat acc tgg aaa atg ctt ctg tcc gtt    144
328 Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Met Leu Leu Ser Val
329           35                  40                  45
331 tgc cgg tcg tgg gcg gca tgg tgc aag ttg aat aac cgg aaa tgg ttt    192
332 Cys Arg Ser Trp Ala Ala Trp Cys Lys Leu Asn Asn Arg Lys Trp Phe
333       50                  55                  60
335 ccc gca gaa cct gaa gat gtt cgc gat tat ctt cta tat ctt cag gcg    240
336 Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu Tyr Leu Gln Ala
337   65                  70                  75                  80
339 cgc ggt ctg gca gta aaa act atc cag caa cat ttg ggc cag cta aac    288
340 Arg Gly Leu Ala Val Lys Thr Ile Gln Gln His Leu Gly Gln Leu Asn
341               85                  90                  95
```

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/853,033

DATE: 08/14/2001
TIME: 12:13:53

Input Set : A:\65691222.app
Output Set: N:\CRF3\08142001\I853033.raw